# Supporting Information

Supporting Information Corrected January 23, 2013

Zhu et al. 10.1073/pnas.1218509110

## SI Materials and Methods

**Cloning, expression and purification of the H17 HA proteins.** The ectodomain (residues 30-527, equivalent to 11-329 of HA1 and 1-174 of HA2 in H3 numbering) of HA from bat influenza virus GU10-060 (H17N10, GenBank accession number CY103892) was expressed in a baculovirus expression system. The ectodomain coding region was inserted into a baculovirus transfer vector, pFastbacHT-A (Invitrogen) with an N-terminal gp67 signal peptide, a C-terminal thrombin cleavage site, a foldon trimerization sequence and a $His_6$-tag, and expressed as described previously (1). The HA HA2-A47G mutation was introduced by site-directed mutagenesis of the GU10-060 ectodomain plasmid. Briefly, the constructed plasmids were used to transform DH10bac competent bacterial cells by site-specific transposition (Tn-7 mediated) to form a recombinant bacmid with beta-galactosidase blue-white receptor selection. The purified recombinant bacmids were used to transfect Sf9 insect cells for overexpression. The HA proteins were produced by infecting suspension cultures of Hi5 cells with recombinant baculovirus at an MOI of 5-10 and incubated at 28ºC shaking at 110 RPM. After 72 hours, Hi5 cells were removed by centrifugation and supernatants containing secreted, soluble HAs were concentrated and buffer-exchanged into 20 mM Tris pH 8.0, 300 mM NaCl, and further purified by metal affinity chromatography using Ni-nitrilotriacetic acid (NTA) resin (Qiagen) to be used for glycan microarray evaluation. For crystal structure determination, the HA mutant was digested with thrombin to remove the foldon domain and $His_6$-tag. The cleaved trimeric H17 HA mutant was

purified further by size exclusion chromatography on a Hiload 16/90 Superdex 200 column (GE healthcare) in 20 mM Tris pH 8.0, 100 mM NaCl and 0.02% (vol/vol) $NaN_3$.

**HA glycan microarray receptor binding assay.** Protocols for microarray HA analysis and the glycan list (Fig. S2) were as previously described (2, 3). Briefly, for analysis with recombinant HA, HA-antibody complexes were prepared by mixing 15 µg of recombinant HA, mouse anti-His Alexa Fluor 488 (Qiagen) and goat anti-mouse IgG Alexa Fluor 488 (Invitrogen) in a molar ratio of 4:2:1, respectively, in 20 mM Tris pH 8.0, 100 mM NaCl, and 0.02% (vol/vol) $NaN_3$ buffer. These prepared complexes were allowed to form for 15 min on ice, and 100 µL of the complex mixture was then added directly to the surface of the array and allowed to incubate in a humidified chamber, protected from the light, for 1 hour at room temperature (~ 22 ºC). Following the initial incubation, HA-antibody solution was removed by pipetting the solution and washing 3 times with 100 µL 1x PBS + 0.05% Tween, pH 7.4, and, subsequently, by dipping 3 times in 1x PBS and then 3 times in distilled $H_2O$. Washed slides were dried by centrifugation and scanned on a ProScanArray Express HT (PerkinElmer) confocal slide scanner for AlexaFluor488 setting. Image data were stored as a TIFF image and signal data was collected using Imagene (BioDiscovery) imaging software. The signal data were processed to determine averaged (mean signal minus mean background) values of 4 replicate spots on the array for each unique printed glycan.

**Crystal structure determination of the H17 HA mutant.** Crystallization experiments were set up using the sitting drop vapor diffusion method. The GU10-060 HA mutant at 10 mg/ml in 20 mM Tris pH 8.0, 100 mM NaCl and 0.02% (vol/vol) $NaN_3$ was crystallized in 0.1 M sodium citrate, pH 5.5, 1 M $LiCl_2$, and 15% (wt/vol) polyethylene glycol (PEG) 6000. The GU10-060 HA crystals were flashed-cooled at 100 K using 20 % ethylene glycol (vol/vol) as a

cryoprotectant. Diffraction data were collected at beamline 08ID-1 at the Canadian Light Source (CLS) (Table 1). Data were integrated and scaled with HKL2000 (4). Data collection statistics are outlined in Table 1.

The GU10-060 HA structure was determined by molecular replacement using the program Phaser (5). The HA structure was determined using the A/California/04/2009 (H1N1) H1 HA structure (PDB ID code 3UBQ). Initial rigid body refinement was performed in Refmac5 (6), and the Refmac-restrained refinement including non-crystallographic symmetry restraints and jelly-body refinement were carried out. Between rounds of refinements, model building was carried out with the program Coot (7). Final statistics for both structures are represented in Table 1. The quality of the structures was analyzed using the JCSG validation suite (www.jcsg.org). All figures were generated with Bobscript (8) except for Figs. 4, S1 and S2, as well as Figs. 3 and S3 which were generated with PyMol (www.pymol.org).

**Protease susceptibility assay**. Protocols for trypsin susceptibility analysis were as previously described (9). For GU10-060 HA2-47G HA, each reaction contained ~5.0 μg of the HA0 and 1% (wt/vol) dodecylmaltoside (to prevent aggregation of the post-fusion HA). 100 mM sodium acetate buffer used for pH 4.9 and pH 4.0, and 20 mM Tris buffer was used for pH 8.0. Reactions were thoroughly mixed, centrifuged at >12,000 g for 30 seconds and allowed to incubate at 37 °C for one hour. After incubation, reactions were equilibrated to room temperature (~22 °C) and the pH was neutralized by addition of 200 mM Tris, pH 8.4. Trypsin was added to all samples except controls, at a final ratio of 1:10 (wt/wt) of trypsin to the HA. Samples were digested overnight (~16 hours) at 22°C. Reactions were quenched by addition of reducing SDS buffer and were boiled for ~2 min. Samples were then analyzed by SDS-PAGE.

**Table S1.** Comparison of GU10-060 HA2-G47 HA with other influenza virus HAs[a].

| Group | Subtype | Strain name | PDB code | HA1 | HA2 | Monomer | RBS subdomain[b] |
|---|---|---|---|---|---|---|---|
| 1 | H1 | A/California/04/2009 | 3UBQ | 2.2[c] (300)[d] | 0.9 (157) | 2.1 (450) | 1.3 (145) |
| 1 | H2 | A/Japan/ 305/1957 | 3KU6 | 2.1 (304) | 1.2 (157) | 2.2 (455) | 1.6 (145) |
| 1 | H5 | A/Vietnam/1203/2004 | 2FK0 | 2.2 (293) | 1.0 (164) | 2.0 (449) | 1.3 (146) |
| 1 | H9 | A/swine/Hong Kong/9/98 | 1JSD | 1.9 (298) | 1.2 (144) | 1.8 (444) | 1.5 (136) |
| 2 | H3 | A/HongKong/19/1968 | 2HMG | 2.3 (293) | 1.4 (161) | 2.4 (452) | 1.5 (141) |
| 2 | H7 | A/turkey/Italy/8458/2002 | 1TI8 | 2.3 (290) | 1.4 (157) | 2.5 (432) | 1.4 (138) |
| 2 | H14 | A/mallard/Astrakhan/263/1982 | 3EYJ | 2.4 (298) | 1.4 (160) | 2.4 (454) | 1.7 (140) |
| | B/HA | B/Hong Kong/8/73 | 3BT6 | 3.0 (257) | 2.8 (140) | 4.0 (390) | 2.5 (141) |

[a] To analyze differences in the overall structure, rmsd values (Å) of $C_\alpha$ atoms were calculated between domains of different HAs superimposed by structural alignment onto the equivalent domain of GU10-060 HA2-G47 HA.

[b] Receptor binding subdomain (RBS subdomain) of HA, corresponding to residues 117-265 of HA1 (H3 numbering).

[c] rmsd values for the pairwise comparison of $C_\alpha$ positions (Å).

[d] Numbers in parentheses represent number of residues that were included in the comparisons.

**Table S2**. Sequence comparison after sequence and structural alignment of HA1 (top) and HA2 (bottom) of GU10-060 HA with GU09-164 HA or other HAs[a].

```
                                      1       10        20        30        40        50  54
GU10-060[H17N10]  -----------MELIILLILLNPYTFVLG----DRICIGYQANQNNQTVNTLLEQNVPV-TGAQEILETNHNGKLCSL
GU09-164[H17N10]  -----------MELIVLLILLNPYTFVLG----DRICIGYQANQNNQTVNTLLEQNVPV-TGAQEILETNHNGKLCSL
CA04[H1N1]        -----------MKAILVVLLYTFATANA----DTLCIGYHANNSTDTVDTVLEKNVTV-THSVNLLEDKHNGKLCKL
SC18 [H1N1]       -----------MEARLLVLLCAFAATNA----DTICIGYHANNSTDTVDTVLEKNVTV-THSVNLLEDSHNGKLCKL
HK68[H3N2]        -------MKTIIALSYIFCLALGQDLPGNDNSTATLCLGHHAVPNGTLVKTITDDQIEV-TNATELVQSSSTGKICNN
B/HK73            -------MKAIIVLLMVVTSN-----------ADRICTGITSSNSPHVVKTATQGEVNVTGVIPLTTTPTKSHFANLK
                 55    60              70        80        90       100       110       120
GU10-060[H17N10]  NGVPPLDL-----QSCTLAGWLLGNPNCDNLLEAEEWSYIKINENAPDDLCFPG-NF-ENLQDLLLEMSGVQNFT-KV
GU09-164[H17N10]  NGVPPLDL-----QSCTLAGWLLGNPNCDSLLEAEEWSYIKINESAPDDLCFPG-NF-ENLQDLLLEMSGVQNFT-KV
CA04[H1N1]        RGVAPLHL-----GKCNIAGWILGNPECESLSTASSWSYIVETPSSDNGTCYPG-DF-IDYEELREQLSSVSSFE-RF
SC18 [H1N1]       KGIAPLQL-----GKCNIAGWLLGNPECDLLLTASSWSYIVETSNSENGTCYPG-DF-IDYEELREQLSSVSSFE-KF
HK68[H3N2]        P-HRILDG-----IDCTLIDALLGDPHCDVFQN-ETWDLFVERSKAFSN-CYPY-DV-PDYASLRSLVASSGTLE-FI
B/HK73            G-TQTRGKLCPNCLNCTDLDVALGRPKCMG-TIPSAKASILHEVKPVTSGCFPIMHDRTKIRQLPNLLRGYENIRLSA
                122           130       140       150       160       170       180       189
GU10-060[H17N10]  KLFN-PQSMT--GVTTN-NVDQTCPFE-GKPSFYRNLNWIQG--NSGLP-FNIEIKNPTS--NPLLLLWGIHNTKDAA
GU09-164[H17N10]  KLFN-PQSMT--GVTTN-NVDQTCPFE-GKPSFYRNLNWIQG--NSGLP-FNIEIKNPTS--NPLLLLWGIHNTKDAA
CA04[H1N1]        EIFPKTSSWP--NHDSNKGVTAACPHA-GAKSFYKNLIWLVKKGNSYPK-LSKSYINDKG--KEVLVLWGIHHPSTSA
SC18 [H1N1]       EIFPKTSSWP--NHETTKGVTAACSYA-GASSFYRNLLWLTKKGSSYPK-LSKSYVNNKG--KEVLVLWGVHHPPTGT
HK68[H3N2]        TEGF---TWT--GVTQN-GGSNACKRG-PGNGFFSRLNWLTKSGSTYPV-LNVTMPNNDN--FDKLYIWGVHHPSTNQ
B/HK73            RNVTNAETAPGGPYIV--GTSGSCPNVTNGNGFFATMAWAVPKNKTATNPLTVEVPYICTKGEDQITVWGFHSD-DET
                190       200       210       220       230       240       250       260
GU10-060[H17N10]  QQRNLYGNDY-SYTIFNFGEKSEEFRPDIGQ-RDE---IKA-HQDRIDYYWGSLPAQSTLRIESTGNLIAPEYGFYYK
GU09-164[H17N10]  QQRNLYGNDY-SYTIFNFGEKSEEFRPEIGQ-RDE---VKA-HQDRIDYYWGSLPAQSTLRIESTGNLIAPEYGFYYK
CA04[H1N1]        DQQSLYQNAD-TYVFVGSSRYSKKFKPEIAI-RPK---VRD-QEGRMNYYWTLVEPGDKITFEATGNLVVPRYAFAME
SC18 [H1N1]       DQQSLYQNAD-AYVSVGSSKYNRRFTPEIAA-RPK---VRD-QAGRMNYYWTLLEPGDTITFEATGNLIAPWYAFALN
HK68[H3N2]        EQTSLYVQES-GRVTVSTRRSQQSIIPNIGS-RPW---VRG-QSSRISIYWTIVKPGDVLVINSNGNLIAPRGYFKMR
B/HK73            QMVKLYGDSKPQKFTSSANGVTTHYVSQIGGFPNQAEDEGLPQSGRIVVDYMVQKPGKTGTIAYQRGVLLPQKVWCAS
                262       270       280       290       300       310       320       329
GU10-060[H17N10]  RKEGKGGLMKSKLPISDCSTKCQTP-LGALNSTL-PFQNVHQQTIGNCPKYVKATSLMLATGL--RNNP--QMEGR
GU09-164[H17N10]  RKEGKGGLMKSKLPISDCSTKCQTP-LGALNSTL-PFQNVHQQTIGNCPKYVKATSLMLATGL--RNNP--QMEGR
CA04[H1N1]        RN-AGSGIIISDTPVHDCNTTCQTP-KGAINTSL-PFQNIHPITIGKCPKYVKSTKLRLATGL--RNIP--SIQSR
SC18 [H1N1]       RG-SGSGIITSDAPVHDCNTKCQTP-HGAINSSL-PFQNIHPVTIGECPKYVRSTKLRMATGL--RNIP--SIQSR
HK68[H3N2]        TG-K-SSIMSSDAPIDTCISECITP-NGSIPNDK-PFQNVNKITYGACPKYVKQNTLKLATGM--RNVP--EKQTR
B/HK73            GR-S-KVIKGSLPLIG-EAD-CLHEKYGGLNKSKPYYTGEHAKAIGNCPIWVKT--PLKLANGTKYRPPAKLLKER

                  1       10        20        30        40        50        60        70      78
GU10-060[H17N10]  GLFGAIAGFIEGGWQGMIDGWYGYHHENQEGSGYAADKEATQKAVDAITNKVNSIIDKMNSQFESNIKEFNRLELRIQ
GU09-164[H17N10]  GLFGAIAGFIEGGWQGMIDGWYGYHHENQEGSGYAADKEATQKAVDAITNKVNSIIDKMNSQFESNIKEFNRLELRIQ
CA04[H1N1]        GLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADLKSTQNAIDEITNKVNSVIEKMNTQFTAVGKEFNHLEKRIE
SC18 [H1N1]       GLFGAIAGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAIDGITNKVNSVIEKMNTQFTAVGKEFNNLERRIE
HK68[H3N2]        GLFGAIAGFIENGWEGMIDGWYGFRHQNSEGTGQAADLKSTQAAIDQINGKLNRVIEKTNEKFHQIEKEFSEVEGRIQ
B/HK73            GFFGAIAGFLEGGWEGMIAGWHGYTSHGAHGVAVAADLKSTQEAINKITKNLNSLSELEVKNLQRLSGAMDELHNEIL
                  80        90        100       110       120       130       140       150   156
GU10-060[H17N10]  HLSDRVDDALLDIWSYNTELLVLLENERTLDFHDANVKNLFEKVKAQLKDNAIDEGNGCFLLLHKCNNSCMDDIKNGT
GU09-164[H17N10]  HLSDRVDDALLDIWSYNTELLVLLENERTLDFHDANVKNLFEKVKAQLKDNAIDEGNGCFLLLHKCNNSCMDDIKNGT
CA04[H1N1]        NLNKKVDDGFLDIWTYNAELLVLLENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCDNTCMESVKNGT
SC18 [H1N1]       NLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSNVRNLYEKVKSQLKNNAKEIGNGCFEFYHKCDDACMESVRNGT
HK68[H3N2]        DLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFEKTRRQLRENAEDMGNGCFKIYHKCDNACIESIRNGT
B/HK73            ELDEKVDDLRADTISSQIELAVLLSNEGIINSEDEHLLALERKLKKMLGPSAVDIGNGCFETKHKCNQTCLDRIAAGT
                   160       170   175
GU10-060[H17N10]  YKYMDYREESHIEK-QKIDG
GU09-164[H17N10]  YKYMDYREESHIEK-QKIDG
CA04[H1N1]        YDYPKYSEEAKLNR-EEIDG
SC18 [H1N1]       YDYPKYSEESKLNR-EEIDG
HK68[H3N2]        YDHDVYRDEALNNR-FQIKG
B/HK73            FNAGEFSLPTFDSLNITAAS
```

[a] Abbreviations: GU10-060, A/little yellow-shouldered bat/Guatemala/060/2010 (H17N10); GU09-164, A/little yellow-shouldered bat/Guatemala/164/2009 (H17N10); CA04, A/California/04/2009 (H1N1); SC18, A/South Carolina/1/18 (H1N1); HK68, A/Hong Kong/68 (H3N2); B/HK73, B/Hong Kong/8/73. Green indicates important residues around the receptor-binding site in all flu A HAs (10). Sequence differences between GU10-060 and GU09-164 HAs are highlighted in cyan. The putative fusion peptide in GU10-060 HA is colored red and highly conserved with all other HAs.

**Table S3**. Conservation of key residues in the receptor-binding site of influenza A HAs [a].

| Amino acid | Residue number | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 98 | 134 | 136 | 153 | 155 | 183 | 190 | 194 | 195 | 225 | 226 | 228 |
| Ala | 0 | 1 | 11 | 0 | 4 | 0 | 367 | 0 | 0 | 4 | 2 | 5 |
| Cys | 4 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| Asp | 1 | 0 | 2 | 0 | 0 | 1 | 6,497 | 0 | 0 | 4514 | 0 | 2 |
| Glu | 0 | 0 | 0 | 0 | 1 | 0 | 5,577 | 0 | 0 | 210 | 0 | 2 |
| Phe | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 0 |
| Gly | 0 | 13,265 | 2 | 0 | 4 | 0 | 19 | 0 | 0 | 7,362 | 0 | 10,482 |
| His | 2 | 0 | 0 | 0 | 878 | 12,788 | 1 | 0 | 5 | 1 | 4 | 0 |
| Ile | 0 | 0 | 1 | 0 | 2,696 | 0 | 2 | 301 | 0 | 0 | 1,185 | 0 |
| Lys | 0 | 0 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 5 | 2 | 0 |
| Leu | 0 | 0 | 0 | 0 | 542 | 10 | 7 | 12,922 | 0 | 0 | 922 | 0 |
| Met | 0 | 0 | 0 | 1 | 0 | 0 | 3 | 0 | 0 | 0 | 5 | 0 |
| Asn | 0 | 2 | 0 | 0 | 0 | 464 | 309 | 0 | 1 | 968 | 0 | 0 |
| Pro | 0 | 0 | 3 | 0 | 1 | 2 | 0 | 16 | 0 | 0 | 3 | 0 |
| Gln | 0 | 0 | 0 | 0 | 2 | 1 | 5 | 0 | 0 | 0 | 10,029 | 0 |
| Arg | 0 | 1 | 1 | 3 | 0 | 3 | 0 | 1 | 0 | 3 | 60 | 9 |
| Ser | 0 | 1 | 8,104 | 1 | 4 | 0 | 1 | 0 | 1 | 6 | 0 | 2,646 |
| Thr | 0 | 0 | 5,150 | 0 | 5,124 | 0 | 121 | 1 | 0 | 3 | 0 | 0 |
| Val | 0 | 2 | 0 | 0 | 3,698 | 0 | 231 | 22 | 0 | 1 | 919 | 1 |
| Trp | 0 | 2 | 0 | 13,267 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Tyr | 13,256 | 0 | 0 | 0 | 264 | 0 | 2 | 0 | 13,260 | 0 | 0 | 0 |
| | | | | | | | | | | | | |
| H17 [b] | Phe [d] | Asn [c] | Asp [c] | Trp | Gln [c] | His | Gln [d] | Leu | Tyr | Ala [d] | His [d] | Asp [c] |

[a] The incidence of an amino acid occurring at certain position is shown. A total of 13,282 full-length, non-redundant HA sequences from all influenza A viruses were available in the Influenza A Virus Resource at the NCBI in September 4, 2012.

[b] The consensus putative receptor-binding site amino acids of two bat GU09-164 and GU10-060 H17 HAs are shown in red numbers in the table. Clearly, sequences at certain positions are unique to bat HAs.

[c] Only present in two bat GU09-164 and GU10-060 H17s.

[d] Present in two bat H17 HAs and only a few other flu A virus HAs.

## Supplementary References


1. Stevens J, *et al.* (2004) Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* 303:1866-1870.
2. Blixt O, *et al.* (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci USA* 101:17033-17038.
3. Zhu X, *et al.* (2012) Influenza virus neuraminidase with reduced enzymatic activity that avidly bind sialic acid receptors. *J Virol* 109:18903-18908.
4. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzymol* 276:307-326.
5. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, & Read RJ (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr D Biol Crystallogr* 61:458-464.
6. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53:240-255.
7. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
8. Esnouf RM (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J Mol Graph Model* 15:132-134, 112-133.
9. Ekiert DC, *et al.* (2009) Antibody recognition of a highly conserved influenza virus epitope. *Science* 324:246-251.
10. Skehel JJ & Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* 69:531-569.
11. Xu R, *et al.* (2012) Functional balance of the hemagglutinin and neuraminidase activities accompanies the emergence of the 2009 H1N1 influenza pandemic. *J Virol* 86:9221-9232.
12. Baker NA, Sept D, Joseph S, Holst MJ, & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037-10041.

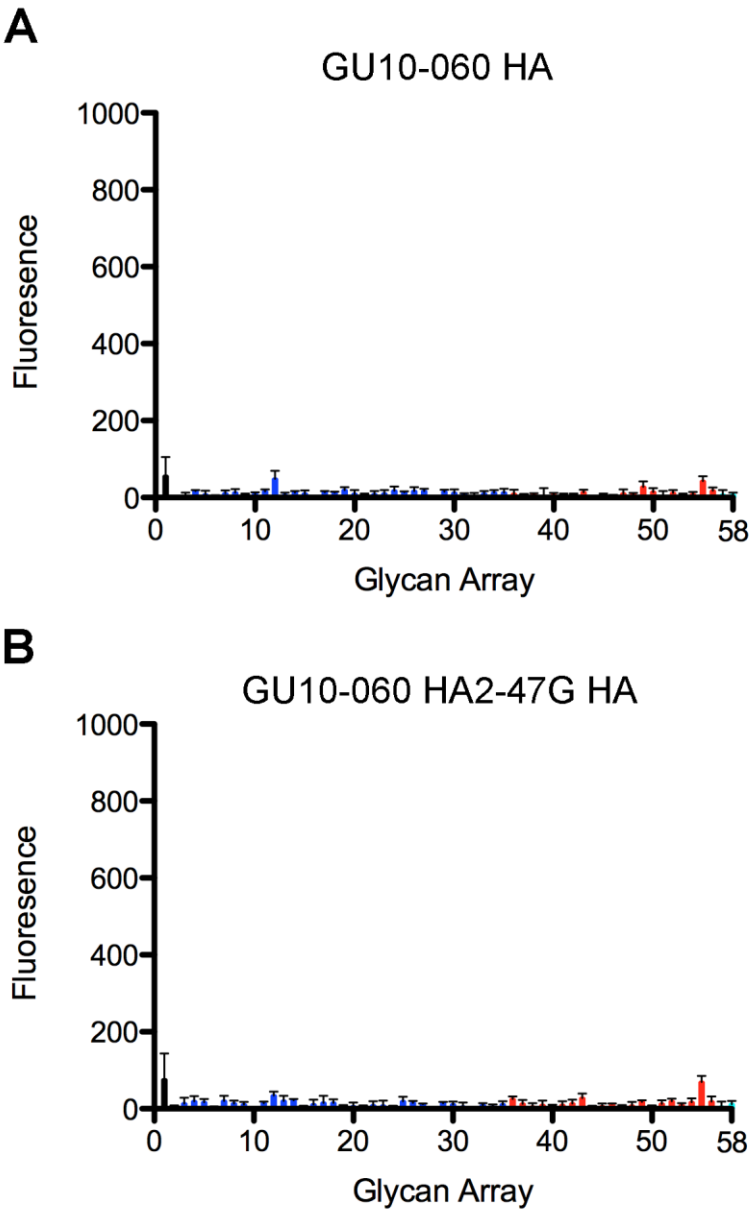
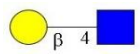
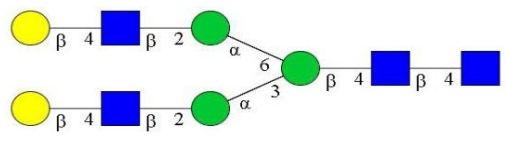
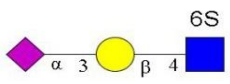
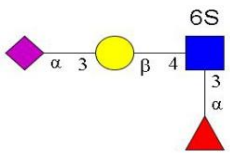
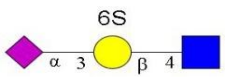
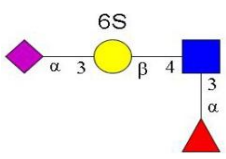
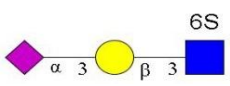
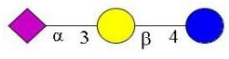
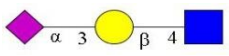
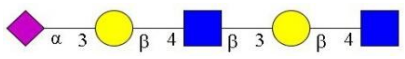
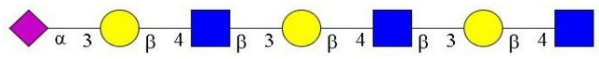
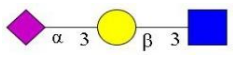
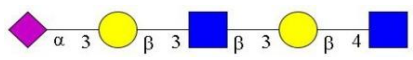
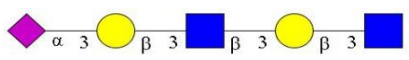
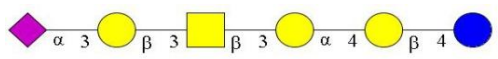
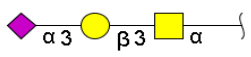
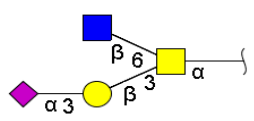
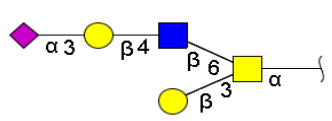
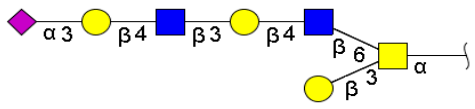
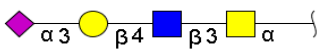
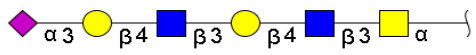
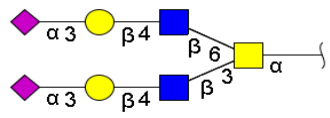
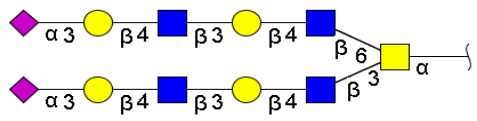
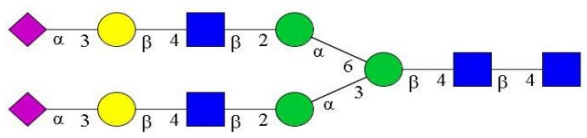
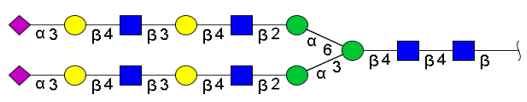
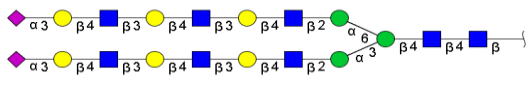
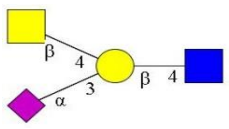
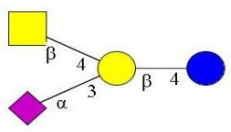
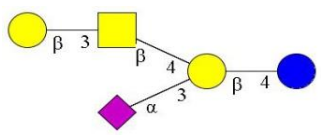
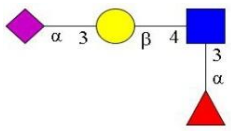
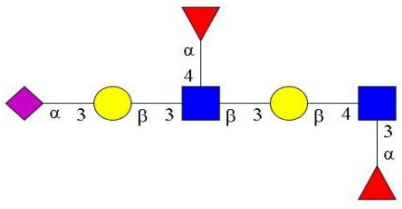


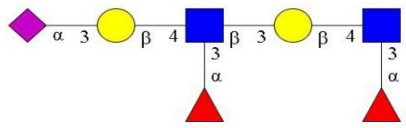
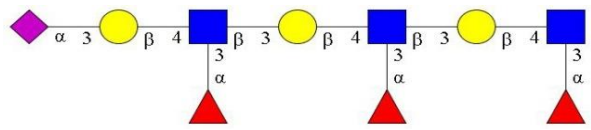
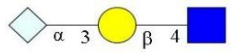
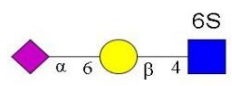
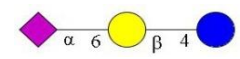
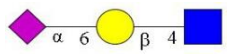
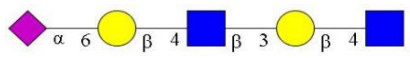
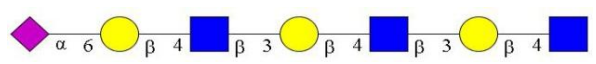
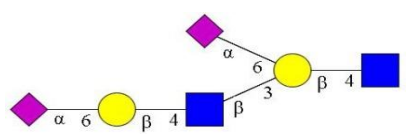
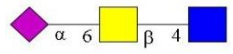
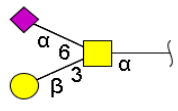

**Fig. S1.** Receptor binding was investigated against printed glycans on a microarray (Fig. S2) for GU10-060 HA (A) and its HA2 A47G mutant (B). The normal starting HA concentration of 15 µg/ml was applied. To observe some background noise, a reduced scale of 1,000 for Y axis (fluorescence signal) was used to make the figure (the typical average non-saturated HA binding signals range from approximately 10,000 to 40,000) (3, 11). Binding signals are shown in filled bars by sialic acid linkages using the same color scheme in Fig. S2. GU10-060 HA and its HA2 A47G mutant exhibit no binding to sialoside glycans on the array.

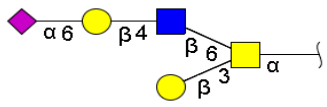
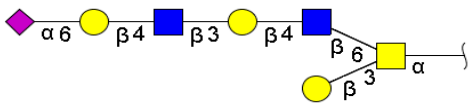
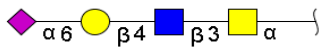
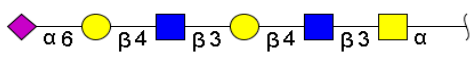
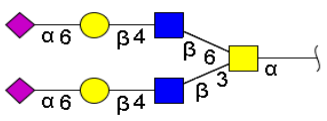
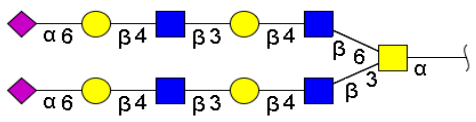
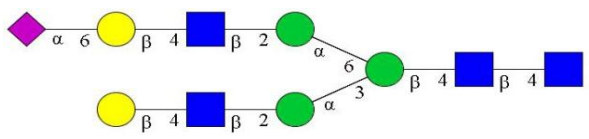
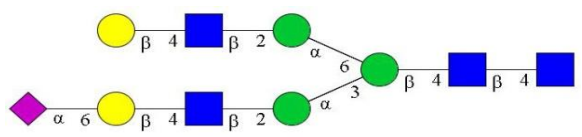
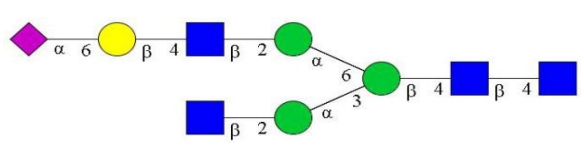
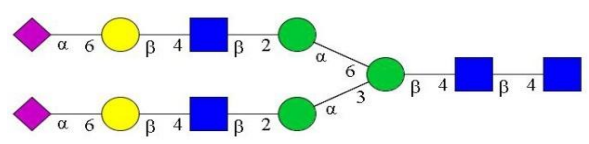
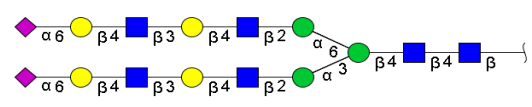
**Fig. S2.** List of glycans on the microarray

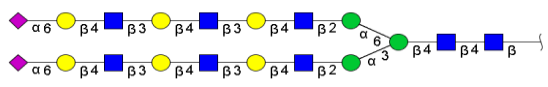
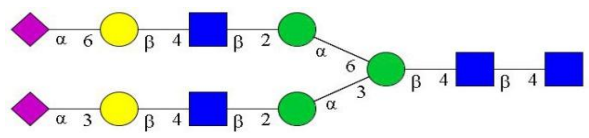
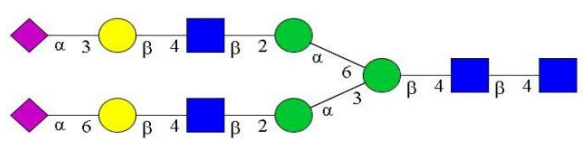
| Chart# | Structure | Name |
|---|---|---|
| 1 |  | Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| 2 |  | Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 3 |  | NeuAcα(2-3)-Galβ(1-4)-6-O-sulfo-GlcNAcβ-propyl-$NH_2$ |
| 4 |  | NeuAcα(2-3)-Galβ(1-4)-[Fucα(1-3)]-6-O-sulfo-GlcNAcβ-propyl-$NH_2$ |
| 5 |  | NeuAcα(2-3)-6-O-sulfo-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| 6 |  | NeuAcα(2-3)-6-O-sulfo-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ-propyl-$NH_2$ |
| 7 |  | NeuAcα(2-3)-Galβ(1-3)-6-O-sulfo-GlcNAcβ-propyl-$NH_2$ |
| 8 |  | NeuAcα(2-3)-Galβ(1-4)-Glcβ-ethyl-$NH_2$ |
| 9 |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| 10 |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |

| | | |
|---|---|---|
| **11** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **12** |  | NeuAcα(2-3)-GalNAcβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **13** |  | NeuAcα(2-3)-Galβ(1-3)-GlcNAcβ-ethyl-$NH_2$ |
| **14** |  | NeuAcα(2-3)-Galβ(1-3)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **15** |  | NeuAcα(2-3)-Galβ(1-3)-GlcNAcβ(1-3)-Galβ(1-3)-GlcNAcβ-ethyl-$NH_2$ |
| **16** |  | NeuAcα(2-3)-Galβ(1-3)-GalNAcβ(1-3)-Gala(1-4)-Galβ(1-4)-Glcβ-ethyl-$NH_2$ |
| **17** |  | NeuAcα(2-3)-Galβ(1-3)-GalNAcα-Thr-$NH_2$ |
| **18** |  | NeuAcα(2-3)-Galβ(1-3)-[GlcNAcβ(1-6)]-GalNAcα-Thr-$NH_2$ |
| **19** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-6)-[Galβ(1-3)]-GalNAcα-Thr-$NH_2$ |
| **20** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-6)-[Galβ(1-3)]-GalNAcα-Thr-$NH_2$ |
| **21** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-GalNAcα-Thr-$NH_2$ |

| | | |
|---|---|---|
| **22** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-GalNAcα-Thr-$NH_2$ |
| **23** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-[NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-6)]-GalNAcα-Thr-$NH_2$ |
| **24** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-[NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-6)]-GalNAcα-Thr-$NH_2$ |
| **25** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| **26** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| **27** |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| **28** |  | NeuAcα(2-3)-[GalNAcβ(1-4)]-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **29** |  | NeuAcα(2-3)-[GalNAcβ(1-4)]-Galβ(1-4)-Glcβ-ethyl-$NH_2$ |
| **30** |  | Galβ(1-3)-GalNAcβ(1-4)-[NeuAcα(2-3)]-Galβ(1-4)-Glcβ-ethyl-$NH_2$ |
| **31** |  | NeuAcα(2-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ-propyl-$NH_2$ |
| **32** |  | NeuAcα(2-3)-Galβ(1-3)-[Fucα(1-4)]-GlcNAcβ(1-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ-ethyl-$NH_2$ |

| | | |
|---|---|---|
| **33** |  | NeuAcα(2-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ(1-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ-ethyl-$NH_2$ |
| **34** |  | NeuAcα(2-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ(1-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ(1-3)-Galβ(1-4)-[Fucα(1-3)]-GlcNAcβ-ethyl-$NH_2$ |
| **35** |  | NeuGcα(2-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **36** |  | NeuAcα(2-6)-Galβ(1-4)-6-O-sulfo-GlcNAcβ-propyl-$NH_2$ |
| **37** |  | NeuAcα(2-6)-Galβ(1-4)-Glcβ-ethyl-$NH_2$ |
| **38** |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **39** |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **40** |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **41** |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-[NeuAcα(2-6)]-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **42** |  | NeuAcα(2-6)-GalNAcβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| **43** |  | NeuAcα(2-6)-[Galβ(1-3)]-GalNAcα-Thr-$NH_2$ |

| | | |
|---|---|---|
| 44 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-6)-[Galβ(1-3)]-GalNAcα-Thr-$NH_2$ |
| 45 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-6)-[Galβ(1-3)]-GalNAcα-Thr-$NH_2$ |
| 46 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-GalNAcα-Thr-$NH_2$ |
| 47 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-GalNAcα-Thr-$NH_2$ |
| 48 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-6)]-GalNAcα-Thr-$NH_2$ |
| 49 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-6)]-GalNAcα-Thr-$NH_2$ |
| 50 |  | Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 51 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 52 |  | GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 53 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 54 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |

| | | |
|---|---|---|
| 55 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 56 |  | NeuGcα(2-6)-Galβ(1-4)-GlcNAcβ-ethyl-$NH_2$ |
| 57 |  | NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |
| 58 |  | NeuAcα(2-6)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-3)-[NeuAcα(2-3)-Galβ(1-4)-GlcNAcβ(1-2)-Manα(1-6)]-Manβ(1-4)-GlcNAcβ(1-4)-GlcNAcβ-Asn-$NH_2$ |

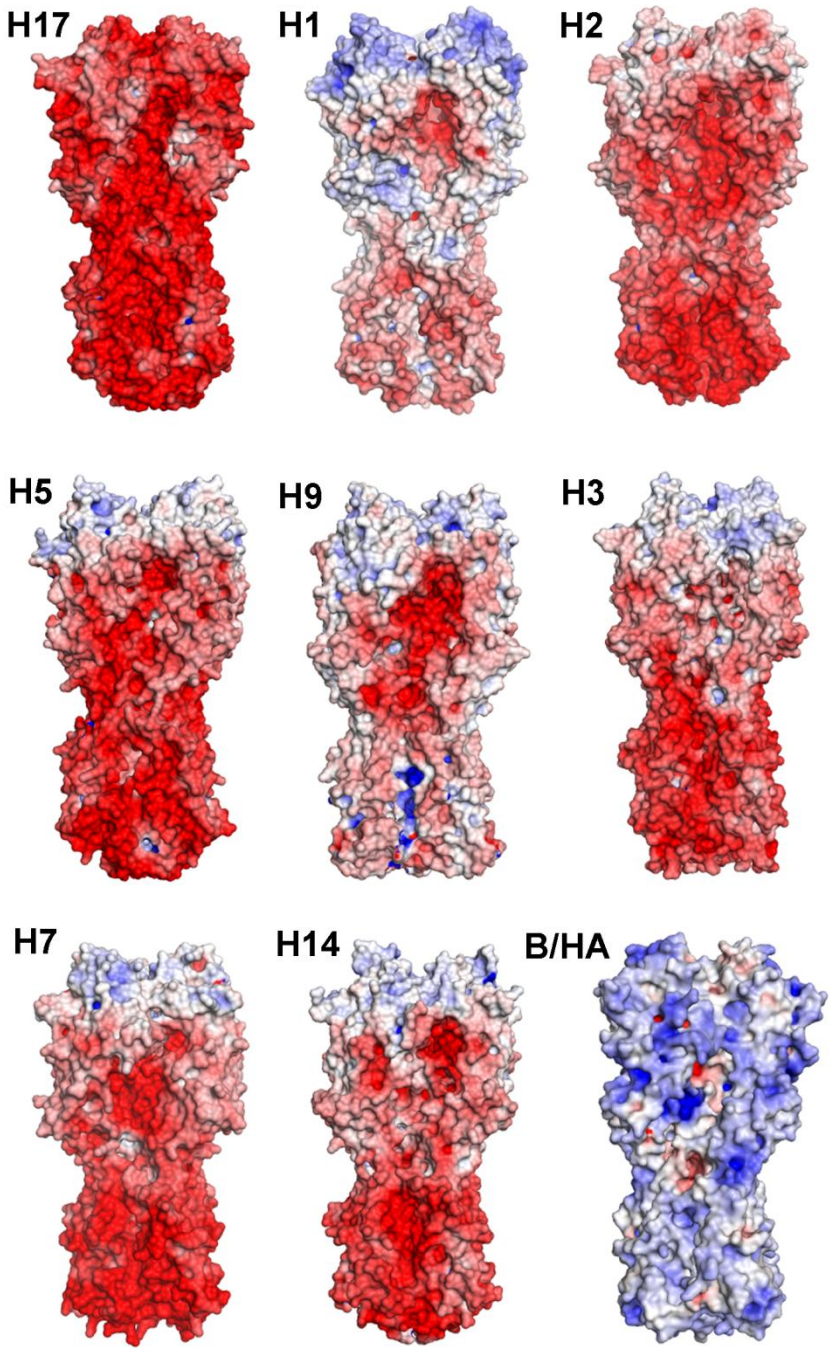



**Fig. S3.** Electrostatic potential surface of the ectodomain trimers of all known influenza virus HAs. Electrostatic surface potentials were calculated using the APBS program (12). Negatively charged regions are red, positively charged regions are blue, and neutral regions are white (-10 to 10 $K_bT/e_c$ potential range). The HA coordinates used in this figure are: flu A group 1 HAs: H17, GU10-060 HA2-47G HA; H1 (PDB ID code 3UBQ), H2 (3KU6), H5 (2FK0), H9 (1JSD); flu A group 2 HAs: H3 (2HMG), H7 (1TI8), H14 (3EYJ) and flu B HA (3BT6). The surface of the H17 HA trimer is unusually acidic.